PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Mini-sternotomy versus conventional sternotomy for aortic valve replacement: a randomised controlled trial
AUTHORS	Hancock, Helen; Maier, Rebecca; Kasim, Adetayo; Mason, James; Murphy, Gavin; Goodwin, Andrew; Owens, W; Akowuah, Enoch

VERSION 1 – REVIEW

REVIEWER	Massimo Chello Università Campus Biomedico di Roma, Rome, Italy
REVIEW RETURNED	03-Sep-2020

GENERAL COMMENTS	Authors should be congratulated for their work.

REVIEWER	Louise Sun University of Ottawa Heart Institute, and Institute for Clinical Evaluative Sciences, Canada
REVIEW RETURNED	07-Sep-2020

GENERAL COMMENTS	Hancock and colleagues conducted a single centre, single-blind RCT of 270 patients, to compare the clinical and economic outcomes of mini sternotomy vs. standard median sternotomy in patients undergoing surgical aortic valve replacement. They did not find a difference between the two techniques in terms of red blood cell transfusion within 7 days. However, they found a lower rate of chest drain losses, longer cardiopulmonary bypass and aortic crossclamp times, and a higher overall cost with the mini sternotomy approach. The authors ought to be commended for eliciting patient feedback while designing the study. Specific Comments: 1. Methods – Statistical Analysis: please specify the period of collection for audit data. Please also detail the method of multiple imputations. 2. Methods – Selection Criteria: did each of the three surgeons have different criteria for determining whether minimally invasive sternotomy is feasible? 3. Methods – Blinding: Is it truly possible for a patient to be blinded to treatment allocation within 48 hours after surgery? They will be able to see their incisions during dressing changes, and may experience localized pain that will inform the size of the incision. 4. Results: cardiopulmonary bypass and aortic crossclamp times are likely right skewed. Could the authors present these data in terms of median (IQR)? Consider presenting descriptive statistics

	for all continuous outcomes as median (IQR), unless normally distributed. 5. A postulated advantage of minimally invasive surgery is faster recovery and decreased narcotic consumption. Have the authors collected data on pain scales and/or narcotic doses? 6. Discussion – Strengths and Limitations: the authors stated that the single center design tended to bias results towards the null. However, the converse may be true (Unverzagt et al. J Clin Epidemiol. 2013;66(11):1271-1280).

REVIEWER	Milan Milojevic Erasmus MC, Netherlands and Dedinje Cardiovascular Institute, Serbia
REVIEW RETURNED	12-Sep-2020

GENERAL COMMENTS	Thank you for the opportunity to review this article by Helen Hancock and colleagues from the UK. The article is well-written, balanced and focused, and the effort to randomize such an extensive number of patients in a single center is exceptional and the authors and all other contributors are to be congratulated on producing such a clear and authorative study that will serve as the landmark for many years in clinical practice guidelines. I have only minor comments: Keywords: - Please include more words such as replacement, cardiac surgery, clinical trial and etc. This may be of help for future meta-analyses. Introduction: - In the first paragraph, I suggest the authors to add some more data from US, Europe and Asia on prevalence of surgical AVR and outcomes (for example, the most recent STS Annual Report from D'Agostino et al in Ann Thorac Surg). - The second paragraph requires some re-writing. The in-hospital outcomes of both surgical and transcatheter AVR is associated with numerous risk factors which are integrated in risk scores (the most famous STS and EuroSCORE). Besides, I suggest the authors to include class I recommendations for both surgical and transcatheter AVR from the 2017 ESC/EACTS guidelines for the management of valvular heart disease to remind the readers about the current evidence-based recommendations from large scales societies. Methods: - I can't understand the author decision to include for the primary endpoint only 7-day time period. This is uncommon measure period for PRBC in cardiac surgery. Usually, number of transfused patients is defined as any blood received during or after surgical procedure until the day of discharge or 30-day, including any blood transfused during a re-operative surgery. I'm aware that there is no room for any change but could you please comment on this? Results: - All-cause death, as the most important outcome measure, is non-discussable and must be part in the main text. In the present form, the results are only presented in table but not in the text. I suggest adding mortality to 'Adverse events' Section to provide more comprehensive details on the difference in hard clinical events. Discussion:

	- Please include the findings of two large RCTs in the discussion section: 1. Quality of life after ministernotomy versus full sternotomy aortic valve replacement doi.org/10.1053/j.semtcvs.2020.07.013 2. The role of ministernotomy in aortic valve surgery-A prospective randomized study doi: 10.1111/jocs.14053. Epub 2019 Apr 24.

REVIEWER	Dr. R. Scott McClure Foothills Medical Centre, University of Calgary Calgary, Alberta Canada
REVIEW RETURNED	13-Sep-2020

GENERAL COMMENTS	The author's have performed a randomized trial comparing manubrium-limited mini-sternotomy to conventional sternotomy for surgical aortic valve replacement. The primary outcome being assess was red-cell transfusion from entrance to the CICU to 7 days postoperatively. The study is well designed and the author's are to be congratulated for their rigour. A consultant driven study, where the 3 consultants have expertise in both surgical techniques being assessed, with efforts to blind patients to the procedure of randomization, with definitive protocol driven transfusion protocols in place. There was minimal cross-over and appropriate intention-to-treat statistical analysis. The economic analysis is an added outcome of interest giving strength to the study. A general comment to varied statistical methods is mentioned to adjust for missing data in the economic analysis and I see bootstrapping techniques have also been used. A formal statistical review by a statistician would be appropriate prior to publication. Overall, the is a well done study that adds to medical literature. I do have some questions/comments to relay. 1. In the conclusion the author's state, "MAVRIC, the world's largest RCT at low risk of bias, found no additional clinical benefit of minimally invasive AVR". The statement "no additional clinical benefit" might be contested. Although not the primary outcome, there was a statistically significant difference in "non-red cell" transfusions. Is the transfusion of any blood product at a statistically higher level something of concern? This can be debated. On the other side of things, lung function was surprisingly worse on POD#4 in the mini-sternotomy group and wound infections were also higher at 12 weeks relative to conventional sternotomy. So maybe these opposing clinical differences cancel each other out - maybe not. My point is, the statement is over reaching. The study found no difference in red-blood cell transfusion, to suggest no additional clinical benefit might be viewed as over reaching. 2. Although a properly powered RCT at its outset, the event rate was substantially less than what was anticipated. Although highly unlikely, (as the event rate is identical across both groups at 23 show no trend in either direction), it could be argued that the study remains underpowered. Technically, this should probably be mentioned in the limitations (I don't feel strongly on this but technically it would be proper).

	3. Presenting of the valve sizes for study - comparing the mean valve size is not a useful parameter when assessing valve implantation The author's should be required to list the distribution of the various valve sizes implanted for the two groups. Then, a statistical comparison should be displayed comparing the number of smaller valves (19mm and 21mm) in one group relative to the number of smaller valves (19mm and 21mm) in the other group. A similar statistical comparison can be performed for larger valves (23mm and up) across the two cohorts. This gives the reader a true assessment of whether smaller or larger valves were inserted in one group over the other. 4. 6 week moderate/severe aortic insufficiency? - In the post-operative valve function paragraph, it notes that 6/134 (mini) and 3/130 (conventional) had moderate to severe aortic insufficiency at follow-up echo assessment. This is quite high and warrants some further explanation. Are these paravalvular leaks? This is what the reader is left to assume but it is not clear. In the Partner 3 trial (NEJM 2018), for the arm of 454 surgical AVR patients (deemed of low risk by STS score) there was a paravalvular leak rate of moderate or severe of 0% at 4 weeks and 0.5% at 1 year. If indeed, in this study, the paravalvular leak rate was 4.5% (6/134) and 2.3% (3/130), a comment to why this occurred is necessary. How many patients left the OR with documented paravalvular leaks? Aside from the above minor constructive suggestions, I feel this to be a very well done study. I congratulate you and your team. Well done.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Authors should be congratulated for their work.

Reviewer: 2
The authors ought to be commended for eliciting patient feedback while designing the study.
Specific Comments:
1. Methods – Statistical Analysis: please specify the period of collection for audit data. Please also detail the method of multiple imputations.

The audit was conducted over 5 years, ending in 2009; this detail has been added to the manuscript.

Details about the method of multiple imputations have also been added to the manuscript.

2. Methods – Selection Criteria: did each of the three surgeons have different criteria for determining whether minimally invasive sternotomy is feasible?

Patients were excluded if only conventional median sternotomy was indicated, for example in the presence of significant skeletal abnormalities like kyphosis. They were also excluded if transoesophageal echocardiography could not be performed, as this was mandatory to perform safe peripheral venous cannulation. All 3 surgeons used consistent criteria. This information has been added to the supplementary appendix.

3. Methods – Blinding: Is it truly possible for a patient to be blinded to treatment allocation within 48 hours after surgery? They will be able to see their incisions during dressing changes, and may experience localized pain that will inform the size of the incision.

Blinding was strictly adhered to in 258 of 270 patients. No dressing changes were allowed during the blinding period. Of the 258 patients who were effectively blinded, only 119 (46%) were able to correctly identify the procedure they had undergone.

4. Results: cardiopulmonary bypass and aortic cross clamp times are likely right skewed. Could the authors present these data in terms of median (IQR)? Consider presenting descriptive statistics for all continuous outcomes as median (IQR), unless normally distributed.

We agree with the reviewer that skewed data should be summarised using median(IQR). However, it is important to consider the severity of the skewness before replacing mean(SD) by median(IQR) because the hypothesis reported was tested on mean difference. The median(IQR) for bypass time were 79(28) and 57(13) for Mini-sternotomy and Conventional sternotomy, respectively. The median (IQR) for Aortic cross clamp time were 62(20) and 44(11) for Mini-sternotomy and Conventional sternotomy, respectively. Although the data are slightly skewed, we do not consider it inappropriate to present mean(SD) for these data given that mean(SD) and median(IQR) are comparable. Given the trade-off, our preference would be that the summary statistics presented in the table are consistent with the measure of effects, i.e. mean difference.

5. A postulated advantage of minimally invasive surgery is faster recovery and decreased narcotic consumption. Have the authors collected data on pain scales and/or narcotic doses?

Data on analgesic use are presented in Table 4 of the supplementary table. Data on pain scores have been added to this table in the revised supplementary appendix.

6. Discussion – Strengths and Limitations: the authors stated that the single center design tended to bias results towards the null. However, the converse may be true (Unverzagt et al. J Clin Epidemiol. 2013;66(11):1271-1280).

We agree with the reviewer that a single centre design can be biased towards either null or alternative hypothesis. However, we believe in the specific case of the MAVRIC trial, that it was biased towards the null because the care protocol administered in this specific centre set the needs for RBC transfusion at a high threshold compared with other NHS Trusts.

Reviewer: 3

The article is well-written, balanced and focused, and the effort to randomize such an extensive number of patients in a single center is exceptional and the authors and all other contributors are to be congratulated on producing such a clear and authorative study that will serve as the landmark for many years in clinical practice guidelines. I have only minor comments:

Keywords: Please include more words such as replacement, cardiac surgery, clinical trial and etc. This may be of help for future meta-analyses.

We have added more key words to the revised manuscript.

Introduction: In the first paragraph, I suggest the authors to add some more data from US, Europe and Asia on prevalence of surgical AVR and outcomes (for example, the most recent STS Annual Report from D'Agostino et al in Ann Thorac Surg).

This has been added to the manuscript. Data from the UK published by National Institute for Cardiac Outcome Reporting has been included to further the point.

The second paragraph requires some re-writing. The in-hospital outcomes of both surgical and transcatheter AVR is associated with numerous risk factors which are integrated in risk scores (the most famous STS and EuroSCORE). Besides, I suggest the authors to include class I recommendations for both surgical and transcatheter AVR from the 2017 ESC/EACTS guidelines for the management of valvular heart disease to remind the readers about the current evidence-based recommendations from large scales societies.

The guidelines above are now included in this part of the manuscript, as are statements about the outcomes depending on Euroscore 2.

Methods: I can't understand the author decision to include for the primary endpoint only 7-day time period. This is uncommon measure period for PRBC in cardiac surgery. Usually, number of transfused patients is defined as any blood received during or after surgical procedure until the day of discharge or 30-day, including any blood transfused during a re-operative surgery. I'm aware that there is no room for any change but could you please comment on this?

Blood transfusion rates including during surgery until discharge are presented in Table 2 and 3 in the main paper. This includes details of red cell and non-red cell transfusions, the number of units to 7 days and number of units to discharge. The number of units transfused outside the period during which the primary outcome was measured was small and analysis of those data do not change the central finding of the trial.

Results: All-cause death, as the most important outcome measure, is non-discussable and must be part in the main text. In the present form, the results are only presented in table but not in the text. I

suggest adding mortality to 'Adverse events' Section to provide more comprehensive details on the difference in hard clinical events.

This detail has been added to the manuscript including in the introduction and results.

Discussion: Please include the findings of two large RCTs in the discussion section:
1.Quality of life after ministernotomy versus full sternotomy aortic valve replacement doi.org/10.1053/j.semtcvs.2020.07.013
2.The role of ministernotomy in aortic valve surgery-A prospective randomized study doi: 10.1111/jocs.14053. Epub 2019 Apr 24.

Both of these trials have been published since the closure of the MAVRIC trial. Both were small, 100 patient trials; their findings have been added to the discussion.

Reviewer: 4

Overall, this a well done study that adds to medical literature. I do have some questions/comments to relay.

1. In the conclusion the author's state, "MAVRIC, the world's largest RCT at low risk of bias, found no additional clinical benefit of minimally invasive AVR". The statement "no additional clinical benefit" might be contested. Although not the primary outcome, there was a statistically significant difference in "non-red cell" transfusions. Is the transfusion of any blood product at a statistically higher level something of concern? This can be debated. On the other side of things, lung function was surprisingly worse on POD#4 in the mini-sternotomy group and wound infections were also higher at 12 weeks relative to conventional sternotomy. So maybe these opposing clinical differences cancel each other out - maybe not. My point is, the statement is over reaching. The study found no difference in red-blood cell transfusion, to suggest no additional clinical benefit might be viewed as over reaching.

The conclusion in the revised manuscript has been modified to reflect that the lack of additional benefit pertained only to blood transfusion rates.

2. Although a properly powered RCT at its outset, the event rate was substantially less than what was anticipated. Although highly unlikely, (as the event rate is identical across both groups at 23 show no trend in either direction), it could be argued that the study remains underpowered. Technically, this should probably be mentioned in the limitations (I don't feel strongly on this but technically it would be proper).

Technically, a trial may be considered underpowered if the values used for power calculation were much different from the actual values from the data. However, this was not the case for this trial. The red cell transfusion rate was assumed to be 13% for the conventional sternotomy group, whilst the actual rate was 17%. The main parameter was the MCID (17%), but this was largely clinically informed. The study

was appropriately powered given the MCID and no power calculation would have detected the observed 0% difference in red cell transfusion rate between the groups.

3. Presenting of the valve sizes for study - comparing the mean valve size is not a useful parameter when assessing valve implantation. The author's should be required to list the distribution of the various valve sizes implanted for the two groups. Then, a statistical comparison should be displayed comparing the number of smaller valves (19mm and 21mm) in one group relative to the number of smaller valves (19mm and 21mm) in the other group. A similar statistical comparison can be performed for larger valves (23mm and up) across the two cohorts. This gives the reader a true assessment of whether smaller or larger valves were inserted in one group over the other.

Details about the median and range of valve sizes implanted in each group are already included in the paper (Table 4).

4. *6 week moderate/severe aortic insufficiency? - In the post-operative valve function paragraph, it notes that 6/134 (mini) and 3/130 (conventional) had moderate to severe aortic insufficiency at follow-up echo assessment. This is quite high and warrants some further explanation. Are these paravalvular leaks? This is what the reader is left to assume but it is not clear.*
In the Partner 3 trial (NEJM 2018), for the arm of 454 surgical AVR patients (deemed of low risk by STS score) there was a paravalvular leak rate of moderate or severe of 0% at 4 weeks and 0.5% at 1 year.
If indeed, in this study, the paravalvular leak rate was 4.5% (6/134) and 2.3% (3/130), a comment to why this occurred is necessary. How many patients left the OR with documented paravalvular leaks?

Only 2 patients in the trial 1 in each arm suffered a paravalvular leak. Both were severe. 7 further patients had moderate regurgitation. These were all intravalvular leaks. Transoesophageal echo was performed in all patients prior to leaving the OR.

VERSION 2 – REVIEW

REVIEWER	Louise Sun, MD, SM, FRCPC, FAHA University of Ottawa Heart Institute, Canada Institute for Clinical Evaluative Sciences, Canada
REVIEW RETURNED	01-Nov-2020

GENERAL COMMENTS	Thank you for addressing my comments. I am satisfied with the revisions.

REVIEWER	Milan Milojevic Dedinje Cardiovascular Institute, Belgrade Serbia, and Erasmus University Medical Center, Rotterdam, The Netherlands
REVIEW RETURNED	01-Nov-2020

GENERAL COMMENTS	Thank you for responding to my comments accordingly. I have no further suggestions for revision. The authors need to be

	congratulated for providing precious information to the academic community.

REVIEWER	Dr. Scott McClure University of Calgary Canada
REVIEW RETURNED	12-Nov-2020

GENERAL COMMENTS	I again congratulate the author's on an excellent study. I feel this to be a well thought out and well done study that will contribute to the literature. Despite my strong support of the author's efforts, they have unfortunately opted not to address a few of my major concerns. I previously mentioned that merely presenting the range (19 - 31) and the median or mean is NOT sufficient when discussing valve size. The hold to this point. The author's suggest they have displayed enough data in Table 4. I respectfully disagree. When comparing a less invasive exposure to that of a more invasive exposure, it is plausible that smaller valves were often used more with one approach than another. And just because the median is the same - it may not tell the whole story. It is most appropriate and again my strong recommendation, that the editor's insist the distribution of valves be presented. The specifics on how to do this I suggested in my prior comments. Second - 7 patients are said to have had "moderate" intravalvular aortic insufficiency after implantation at 6 weeks? Again - this needs some sort of comment. I am not particular on what is said - but it needs to be addressed. What valves were these? Were they sutureless? Did you over balloon dilate? To have moderate AI at 6 weeks, be it paravalvular or intravalvular is a bit odd. And the 2 severe paravalvular leaks also should be discussed. What type of valve? sutured? sutureless? Please address the above 2 constructive suggestions to strengthen your manuscript and to bring improved clarity to the readers. Overall, I again applaud you. I think this to be an excellent study. Well done.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2

I am satisfied with the revisions.

Reviewer: 3

Thank you for responding to my comments accordingly. I have no further suggestions for revision. The authors need to be congratulated for providing precious information to the academic community.

Reviewer: 4

I again congratulate the author's on an excellent study. I feel this to be a well thought out and well done study that will contribute to the literature. Despite my strong support of the author's efforts, they have unfortunately opted not to address a few of my major concerns.

I previously mentioned that merely presenting the range (19 - 31) and the median or mean is NOT sufficient when discussing valve size. The hold to this point. The author's suggest they have displayed enough data in Table 4. I respectfully disagree. When comparing a less invasive exposure to that of a more invasive exposure, it is plausible that smaller valves were often used more with one approach than another. And just because the median is the same - it may not tell the whole story. It is most appropriate and again my strong recommendation, that the editor's insist the distribution of valves be presented. The specifics on how to do this I suggested in my prior comments.

Second - 7 patients are said to have had "moderate" intravalvular aortic insufficiency after implantation at 6 weeks? Again - this needs some sort of comment. I am not particular on what is said - but it needs to be addressed. What valves were these? Were they sutureless? Did you over balloon dilate? To have moderate AI at 6 weeks, be it paravalvular or intravalvular is a bit odd. And the 2 severe paravalvular leaks also should be discussed. What type of valve? sutured? sutureless?

Please address the above 2 constructive suggestions to strengthen your manuscript and to bring improved clarity to the readers. Overall, I again applaud you. I think this to be an excellent study.

Well done.

We have now added additional detail about valve size and other characteristics, to the manuscript.

The submission now includes the information below, alongside the median and range of valve sizes and other valve characteristics by group, which were already included. We have created a new table, Table 5, which now includes all details of valve characteristics, by group, and includes the additional data in the table below. The two figures, also below, have been added as well.

We have also added the following text to the manuscript:

Only 2 patients in the trial, 1 in each arm, suffered a paravalvular leak; both were severe. One of these patients, in the mini sternotomy arm had a sutureless valve prosthesis. 7 further patients had moderate regurgitation; these were all intravalvular leaks. Transoesophageal echo was performed in all patients prior to leaving the operating theatre.

Valve Characteristics	Mini-sternotomy group (n=135)	Conventional sternotomy group (n=135)
Valve size: mm		
19mm - 21mm n (%)	40 (29.6)	38 (28.1)
23mm - 25mm n (%)	84 (62.2)	80 (59.3)
27mm - 31mm n (%)	11 (8.2)	17 (12.6)

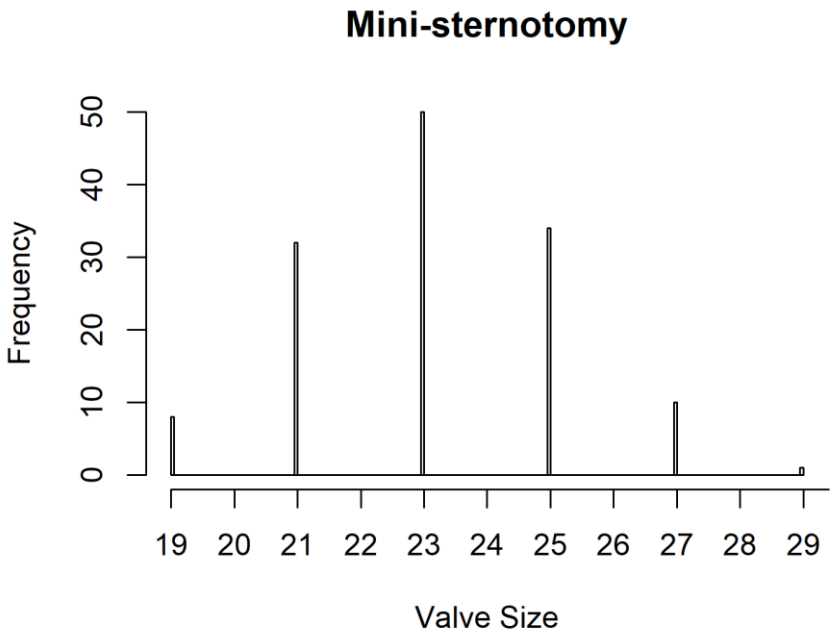

Figure 2. Valve size distribution: mini-sternotomy group

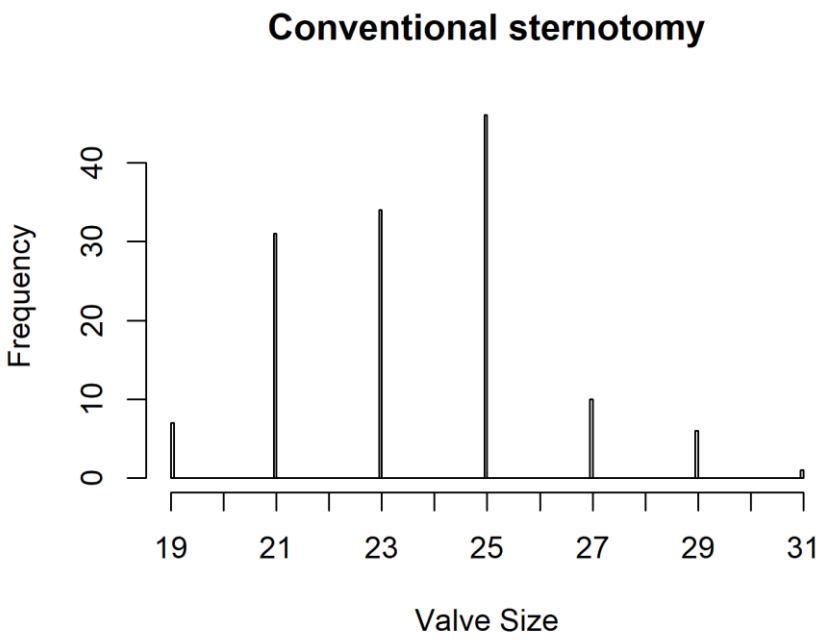

Figure 3. Valve size distribution: conventional sternotomy group

VERSION 3 – REVIEW

REVIEWER	Dr. R. Scott McClure University of Calgary Foothills Medical Centre Libin Cardiovascular Institute Canada
REVIEW RETURNED	05-Dec-2020

GENERAL COMMENTS	The author's have made the appropriate efforts to address my concerns. Their data is transparent and clear. Well done.